Predicting 30-day Readmission risk for ICU patients using a multi-task deep learning framework

By
Aditya Tushar Wadnerkar
Rimzim Thube
Shree Gowri Radhakrishna
Vijaylaxmi Nagurkar

Abstract

Unsupervised clustering of patient sub-populations for multitask learning using electronic health records

Readmission whether planned or unplanned costs the U.S healthcare system billions of dollars annually. This, among other things, is of great importance to healthcare systems, since about one-third of readmissions are preventable, providing a huge opportunity for hospital systems to improve their quality of care (1). This compounded with the fact that hospitals are also penalized with up to 3% lower reimbursements from the Centers for Medicare & Medicaid Services(CMS) for increased readmission rates, as a part of the 2012 affordable care act passed by the Obama administration (2), creates a huge opportunity for the use of data to build algorithms that are capable of predicting which patients are likely to be readmitted if discharged early.

Proposed study

Single task and multi-task models to predict 30-day readmission risk using the MIMIC III dataset. This will be done by first clustering patients using a data-driven unsupervised approach to cluster patients using sociodemographic, operational, and clinical factors captured in the last 48 hours of their stay. This will then be used to predict the patient's risk of readmission within 30 days of discharge from the ICU using a multi-task framework.

Data Source

This project utilized the MIMIC III electronic health record (EHR) dataset, which is comprised of over 58,000 hospital admissions for 38,645 adults and 7,875 neonates. This dataset is a collection of de-identified intensive care unit stays at the Beth Israel Deaconess Medical Center from June 2001- October 2012.

Input

Sociodemographic, operational and expertly crafted disease-specific clinical variables(comorbidities), with a small portion including the time-varying measurements such as labs and vitals to comprehensively capture a patient's progress during their stay.

Output

How many clusters (components)?

Using the embeddings created by the auto-encoder, the Akaike information criterion (AIC) or Bayesian information criterion (BIC) plot will be created to determine the optimal number of components (clusters) that minimize the AIC and BIC.

Roles and Responsibilities

Literature review	Vijaylaxmi/Gauri
Data Preprocessing	Aditya/Rimzim
Architecture study	Aditya/gauri
Model training	Vijaylaxmi/Rimzim
Visualization	Team

References

https://towardsdatascience.com/predicting-30-readmission-risk-for-icu-patients-be3eec7e6681